

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 18:27:17 ; Search time 26.5 Seconds  
(without alignments)  
107.116 Million cell updates/sec

Title: US-09-743-225-4

Perfect score: 62

Sequence: 1 KDKATGTHDG 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	90.3	345	6 Q95LB0	Q95LB0 pan troglod
2	46	74.2	345	11 Q9C142	Q9C142 mus musculu
3	43	69.4	256	2 Q8GJ1	Q8GJ1 lactobacill
4	42	67.7	574	4 Q19301	Q19301 caenorhabdi
5	42	67.7	1674	10 Q64923	Q64923 zea mays (m
6	41	66.1	1397	10 Q9LP90	Q9LP90 arabidopsis
7	40	64.5	1338	10 Q8RUU3	Q8RUU3 oryza sativ
8	40	64.5	1750	10 Q9LPC5	Q9LPC5 arabidopsis
9	39	62.9	391	17 Q9HPW1	Q9HPW1 halobacteri
10	38	61.3	25	5 Q9NBE2	Q9NBE2 chironomus
11	38	61.3	283	3 Q9UOX3	Q9UOX3 coriolus ve
12	38	61.3	606	16 Q9WKN8	Q9WKN8 thermotoga
13	38	61.3	785	3 Q9P8H5	Q9P8H5 humicola in
14	37	59.7	25	5 Q9NBE0	Q9NBE0 chironomus
15	37	59.7	98	5 Q9Y119	Q9Y119 chironomus
16	37	59.7	98	5 Q9Y118	Q9Y118 chironomus

17	37	59.7	98	5 Q9XTN1	Q9XTN1 chironomus
18	37	59.7	145	2 Q8VTC1	Q8VTC1 heliobacte
19	37	59.7	160	5 Q94443	Q94443 chironomus
20	37	59.7	161	5 Q94445	Q94445 chironomus
21	37	59.7	161	5 Q94444	Q94444 chironomus
22	37	59.7	163	5 Q94442	Q94442 chironomus
23	37	59.7	216	16 Q97175	Q97175 clostridium
24	37	59.7	301	2 Q8KKA4	Q8KKA4 proteus vul
25	37	59.7	372	17 Q8ZUR8	Q8ZUR8 pyrobaculum
26	37	59.7	536	16 Q8FWC2	Q8FWC2 brucella su
27	37	59.7	537	16 Q99XX7	Q99XX7 streptococ
28	37	59.7	537	16 Q8NZA9	Q8NZA9 streptococ
29	37	59.7	537	16 Q8KSS1	Q8KSS1 streptococ
30	37	59.7	551	16 Q8YBZ7	Q8YBZ7 brucella me
31	36	58.1	143	16 Q8XVR3	Q8XVR3 ralatonia s
32	36	58.1	307	16 Q9KU39	Q9KU39 vibrio chol
33	36	58.1	376	16 Q92D76	Q92D76 listeria in
34	36	58.1	466	2 Q9S305	Q9S305 ruminococcu
35	36	58.1	591	11 Q8CFS1	Q8CFS1 mus musculu
36	36	58.1	626	17 Q8ZV30	Q8ZV30 pyrobaculum
37	36	58.1	777	10 Q8LP22	Q8LP22 oryza sativ
38	36	58.1	863	11 Q9JMD1	Q9JMD1 mus musculu
39	36	58.1	1437	5 Q9G051	Q9G051 dictyostell
40	35	56.5	161	5 Q9NB94	Q9NB94 anopheles s
41	35	56.5	208	12 Q9DDU0	Q9DDU0 pelargonium
42	35	56.5	232	1 Q9HH39	Q9HH39 methanobact
43	35	56.5	232	9 Q80214	Q80214 methanobact
44	35	56.5	234	16 Q912G9	Q912G9 pseudomonas
45	35	56.5	300	16 Q54392	Q54392 streptomyc

#### ALIGNMENTS

##### RESULT 1

Q95LB0	PRELIMINARY;	PRT;	345 AA.
ID	Q95LB0		
AC	Q95LB0;		
DT	01-DEC-2001 (TREMBlrel. 19, Created)		
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	Apollipoprotein H.		
GN	APOH.		
OS	Pan troglodytes (Chimpanzee).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
OX	NCBI_TaxID=9598;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21372074; PubMed=11479737;		
RA	Sanghera D.K., Nestlerode C.S., Ferrell R.E., Kamboh M.I.;		
RT	"Chimpanzee apolipoprotein H (beta2-glycoprotein I): report on the		
RT	gene structure, a common polymorphism, and a high prevalence of		
RL	antiphospholipid antibodies.";		
RL	Hum. Genet. 109:63-72(2001).		
DR	EMBL; AF358415; AAK71538.1; JOINED.		
DR	EMBL; AF358408; AAK71538.1; JOINED.		
DR	EMBL; AF358409; AAK71538.1; JOINED.		
DR	EMBL; AF358410; AAK71538.1; JOINED.		
DR	EMBL; AF358411; AAK71538.1; JOINED.		
DR	EMBL; AF358412; AAK71538.1; JOINED.		
DR	EMBL; AF358413; AAK71538.1; JOINED.		
DR	EMBL; AF358414; AAK71538.1; JOINED.		
DR	InterPro: IPR000436; Sush1_SCR_CCP.		
DR	Pfam: PF00084; sush1; 4.		
DR	SMART: SM00032; CCP; 4.		
KW	Lipoprotein.		
SQ	SEQUENCE 345 AA; 38267 MW; 712A3EDA2AD4FD36 CRC64;		

Query Match 90.3%; Score 56; DB 6; Length 345;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 KDKATFGTHDG 11
Db 227 KDKATFGCHDG 237

RESULT 2
OQC42 PRELIMINARY; PRT; 345 AA.
AC OQC42; 2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 13 days embryo liver cDNA, RIKEN full-length enriched library,
DE clone:251008B09, full insert sequence.
GN APOH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010926; BAB27271.1;
DR HSP; P02749; IC12.
DR MGD; MGI:88058; Apoh.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 4.
DR SMART; SM00032; CCP; 4.
SQ SEQUENCE 345 AA; 36688 MW; 33242A7E16F1D99B CRC64;

Query Match 74.2%; Score 46; DB 11; Length 345;
Best Local Similarity 80.0%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KDKATFGTHD 10
Db 227 KDKATFGCHE 236

RESULT 3
O8GGJ1 PRELIMINARY; PRT; 256 AA.
AC O8GGJ1; 2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Abc3.
DE Abc3.
GN Abc3.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;

```

```

[1] SEQUENCE FROM N.A.
RP STRAIN=CCM3626;
RA Bringle F., Hubert J.-C.;
RT "Investigation of arginine requirement in Lactobacillus isolated from
RT different environments revealed point mutations, insertions and
RT deletions in arginine biosynthetic genes.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF514870; AA015988.1;
SQ SEQUENCE 256 AA; 28597 MW; B4FE2B25276CEB5C CRC64;

Query Match 69.4%; Score 43; DB 2; Length 256;
Best Local Similarity 63.6%; Pred. No. 5.4;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KDKATFGTHDG 11
Db 242 QDKLNEGAHDG 252

RESULT 4
Q19301 PRELIMINARY; PRT; 574 AA.
AC Q19301;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 62.3 kDa protein.
GN F10E7.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A.;
RT "The sequence of C. elegans cosmid F10E7.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U01264; AAA82423.1;
DR WormPep; F10E7.9; CE04364.
DR InterPro; IPR002293; AA/rel_permease1.
KW Hypothetical protein.
SQ SEQUENCE 574 AA; 62301 MW; AA209D1029FC700E CRC64;

Query Match 67.7%; Score 42; DB 5; Length 574;
Best Local Similarity 72.7%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KDKATFGTHDG 11
Db 42 KDKVLFGTWDG 52

RESULT 5
O64923 PRELIMINARY; PRT; 1674 AA.
AC O64923;
DT 01-AUG-1998 (TREMBlrel. 07, Created)

```

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Starch synthase DULL1.  
 GN DULL1.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98169346; PubMed=9501113;  
 RA Gao M., Wanat J., Stinard P.S., James M.G., Myers A.M.;  
 RT "Characterization of dull1, a maize gene coding for a novel starch  
 RT synthase";  
 RL Plant Cell 10:399-412(1998).  
 DR EMBL; AF023159; AAC14014.1;  
 DR InterPro; IPR001296; Glyco\_transf\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 SQ SEQUENCE 1674 AA; 188311 MW; FA31BE959A01771D CRC64;  
  
 Query Match 67.7%; Score 42; DB 10; Length 1674;  
 Best Local Similarity 72.7%; Pred. No. 65;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 1 KDKATGTHDG 11  
 DB 611 KQLATVGTTHDG 621  
  
 RESULT 6  
 Q9LP90 PRELIMINARY; PRT; 1397 AA.  
 AC Q9LP90;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE T32E20.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC T32E20 from chromosome  
 RT 1";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Lam B.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC020646; AAF79797.1;  
 DR InterPro; IPR000953; Chromo.  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR InterPro; IPR001584; Rve.  
 DR InterPro; IPR000477; RVTse.  
 DR Pfam; PF00385; chromo; 1.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 DR Pfam; PF00665; rve; 1.  
 DR Pfam; PF00078; rvt; 1.  
 DR SMART; SM00298; CHROMO; 1.  
 DR PROSITE; PS50013; CHROMO\_2; 1.  
 KW RNA-directed DNA polymerase; Transferase.  
 SQ SEQUENCE 1397 AA; 159426 MW; IED5717280BF298E CRC64;  
  
 Query Match 66.1%; Score 41; DB 10; Length 1397;  
 Best Local Similarity 70.0%; Pred. No. 82;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 DKATGTHDG 11  
 DB 634 EXTFTTHDG 643  
  
 RESULT 7  
 Q8RUU3 PRELIMINARY; PRT; 1338 AA.  
 AC Q8RUU3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Putative gag-pol polyprotein.  
 GN OJ1004\_F02.14 OR OSJNBA0014J14.7;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Buell C.R., Yuan Q., Qiyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,  
 RA Tsirir T., Riggs F., Heiao J., Zismann V., Blunt S., Pai G.,  
 RA Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 10 BAC OJ1004\_F02 genomic sequence";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
 RA Saski C., Henry D., Oates R., Simmons J.;  
 RT "Rice Genomic Sequence";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC092387; AAL82656.1;  
 DR EMBL; AC092172; AAM18147.1;  
 DR Gramene; Q8RUU3;  
 DR InterPro; IPR000953; Chromo.  
 DR InterPro; IPR005162; Retrotrans\_gag.  
 DR InterPro; IPR001584; Rve.  
 DR InterPro; IPR000477; RVTse.  
 DR Pfam; PF03732; Retrotrans\_gag; 1.  
 DR Pfam; PF00665; rve; 1.  
 DR Pfam; PF00078; rvt; 1.  
 DR SMART; SM00298; CHROMO; 1.  
 KW Polyprotein; RNA-directed DNA polymerase; Transferase.  
 SQ SEQUENCE 1338 AA; 150643 MW; 7D68E2A564279648 CRC64;  
  
 Query Match 64.5%; Score 40; DB 10; Length 1338;  
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 2 DKATGTHDG 11  
 DB 634 EXTFTTHDG 643

Db 513 DKTAFTGHDG 522

RESULT 8

Q9LPC5 PRELIMINARY; PRT; 1750 AA.

AC Q9LPC5

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE F22M8.9 protein.

GN F22M8.9

OS Arabidopsis thaliana (Mouse-ear cross).

OC Spermatophyta; Viridiplantae; Streptophyta; Tracheophyta;

OC Magnoliophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Liu S.X., Sakano H., Yu G., Lee J.M., Lenz C., Pham P., Toriumi M.,

RA Chin C., Chlou J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,

RA Vaysberg M., Alfalfi H., Brooks S., Buehler E., Chao Q., Conn L.,

RA Conway A.B., Hansen N.F., Johnson-Kopson C., Khan S., Kim C., Lam B.,

RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,

RA Ecker J.R., Federspiel N.A., Theologis A.;

RT "The sequence of BAC F22M8 from Arabidopsis thaliana chromosome 1.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC020622; AAF76474.1;

DR HSSP; Q99418; 1PBV.

DR InterPro; IPR006162; Ppantn\_attach.

DR InterPro; IPR000904; Sec7.

DR Pfam; PF01369; Sec7; 1.

DR SMART; SM00222; Sec7; 1.

DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.

DR PROSITE; PS00130; Sec7; 1.

DR SEQUENCE 1750 AA; 194941 MW; F4FCFF663ACDB019 CRC64;

Query Match 64.5%; Score 40; DB 10; Length 1750;

Best Local Similarity 77.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KATFGTHDG 11

Db 264 KGTFGGHDG 272

RESULT 9

Q9HPW1 PRELIMINARY; PRT; 391 AA.

AC Q9HPW1

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE 2',3'-cyclic-nucleotide 2'-phosphodiesterase.

GN Yfkn OR VNG1450G.

OS Halobacterium sp. (strain NRC-1).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

OC Halobacteriaceae; Halobacterium.

OX NCBI\_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE-20504483; PubMed-11016950;

RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,

RA Swartzell S., Weir D., Hall J., Dahl R., Weiti R., Goo Y.A.,

RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.B., Krebs M.P., Angevine C.M., Dale H.,

RA Isenbarger T.A., Peck R.F., Polischroder T., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;

RT "Genome sequence of Halobacterium species NRC-1.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

DR EMBL; AB005061; AAG19756.1;

DR InterPro; IPR006179; 5\_nucleotidase.

DR InterPro; IPR004843; M-peptidase.

DR Pfam; PF02872; 5\_nucleotidase; 1.

DR Pfam; PF00149; Metallophos; 1.

DR PRINTS; PR01607; APYRASEFAMILY.

KW Complete proteome.

SQ SEQUENCE 391 AA; 41347 MW; 03315396A54CE9F5 CRC64;

Query Match 62.9%; Score 39; DB 17; Length 391;

Best Local Similarity 77.8%; Pred. No. 48;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DKATFGTHDG 10

Db 10 DVATFGNHD 18

RESULT 10

Q9NBE2 PRELIMINARY; PRT; 25 AA.

AC Q9NBE2

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Hemoglobin IIB (Fragment).

GN GB2B.

OS Chironomus agilis.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;

OC Chironomidae; Chironominae; Chironomus.

OX NCBI\_TaxID=113487;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE-20336639; PubMed-10876092;

RA Gruhl M.C., Scherbik S.V., Almano K.G., Blinov A., Diez J.-L.,

RA Bergtrom G.;

RT "Insect globin gene polymorphisms: intronic minisatellites and a

RT retroposon interrupting exon 1 of homologous globin genes in

RT Chironomus (Diptera).";

RL Gene 251:153-163(2000).

DR EMBL; AF250302; AAF87711.1;

DR HSSP; P02229; 1ECA.

DR NON\_TER 1

FT NON\_TER 25

FT SEQUENCE 25 AA; 2825 MW; 55BA200610C6F466 CRC64;

Query Match 61.3%; Score 38; DB 5; Length 25;

Best Local Similarity 63.6%; Pred. No. 3.7;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KDKATFGTHDG 11

Db 10 KDTASFATHAG 20

RESULT 11

Q9UOX3 PRELIMINARY; PRT; 283 AA.

AC Q9UOX3

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Heat shock protein 30.

GN CVHSP30/2 OR FDD123B.

OS Coriolus versicolor.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolus.

OX NCBI\_TaxID=57466;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-IFO30340; TISSUE=Mycelium;

RA Iimura Y.;

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

[2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-IFO30340; TISSUE-Myxcelium;  
 RX MEDLINE-97398356; PubMed-9256254;  
 RA Iimura Y., Tatsumi K.;  
 RT "Isolation of mRNAs induced by a hazardous chemical in white-rot  
 fungus, *Coriolus versicolor*, by differential display.";  
 RL FEBS Lett. 412:370-374(1997).  
 DR EMBL; AB018407; BAA76591.1; -  
 DR EMBL; AB018405; BAA76589.1; -  
 DR HSP; P02945; IBRX.  
 DR InterPro; IPR001425; Bac\_rhodopsin.  
 DR Pfam; PF01036; Bac\_rhodopsin; 1.  
 DR PRINTS; PR00251; BACTRLOPSIN.  
 KW Heat shock.  
 SQ SEQUENCE 283 AA; 31031 MW; E9B4DC31CA23F1FA CRC64;

Query Match 61.3%; Score 38; DB 3; Length 283;  
 Best Local Similarity 70.0%; Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DKATFGTHDG 11

Db 243 DYATFGLHSG 252

#### RESULT 12

Q9WXN8 PRELIMINARY; PRT; 606 AA.  
 AC Q9WXN8;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Oligopeptide ABC transporter, periplasmic oligopeptide-binding  
 DE protein.  
 GN TM0031.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=2336;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-MS88 / DSM 3109;  
 RX MEDLINE-99287316; PubMed-10360571;  
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,  
 Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 genome sequence of *Thermotoga maritima*.";  
 RL Nature 399:323-329(1999).  
 DR EMBL; AE001690; AAD35125.1; -  
 DR TIGR; TM0031; -  
 DR InterPro; IPR000914; SBP\_bac\_5.  
 DR Pfam; PF00496; SBP\_bac\_5; 4.  
 KW Complete proteome.  
 SQ SEQUENCE 606 AA; 69941 MW; FC7217EC175A6082 CRC64;

Query Match 61.3%; Score 38; DB 16; Length 606;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KATFGTHDG 11

Db 350 KATFGTHDG 358

#### RESULT 13

Q9P8H5 PRELIMINARY; PRT; 785 AA.  
 ID Q9P8H5;  
 AC Q9P8H5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Cellobiose dehydrogenase.  
 OS Humicola insolens.  
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.  
 OX NCBI\_TaxID=34413;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Golightly E., Brown K.;  
 RT "Nucleic acids encoding polypeptides having cellobiose dehydrogenase  
 activity.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA Xu F., Golightly E.J., Duke K.R., Lassen S.F., Brown K.M., Brown S.H.,  
 Shulein M.;  
 RT "Humicola insolens cellobiose dehydrogenase: cloning, redox chemistry,  
 and 'logic gate'-like dual functionality.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF257654; AAF69005.1; -  
 DR HSP; O01738; ID7C.  
 DR InterPro; IPR000172; GMC\_oxred.  
 DR InterPro; IPR000205; NAD\_binding.  
 DR Pfam; PF00732; GMC\_oxred; 1.  
 DR PROSITE; PS00623; GMC\_OXRED.1; 1.  
 SQ SEQUENCE 785 AA; 84856 MW; 223F5B89A122FBAF CRC64;

Query Match 61.3%; Score 38; DB 3; Length 785;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DKATFGTHD 10

Db 188 DTATFGFHD 196

#### RESULT 14

Q9NBE0 PRELIMINARY; PRT; 25 AA.  
 ID Q9NBE0;  
 AC Q9NBE0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hemoglobin IIB (Fragment).  
 GN G82B.  
 OS Chironomus tepperi.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;  
 OC Chironomidae; Chironominae; Chironomus.  
 OX NCBI\_TaxID=113505;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE-20336639; PubMed-10876092;  
 RA Gruhl M.C., Scherbik S.V., Alamanova K.G., Blinov A., Diez J.-L.,  
 Bergtrom G.;  
 RT "Insect globin gene polymorphisms: intronic minisatellites and a  
 retroposon interrupting exon 1 of homologous globin genes in  
 Chironomus (Diptera).";  
 RL Gene 251:153-163(2000).  
 DR EMBL; AF250304; AAF87713.1; -  
 DR HSP; P02229; 1ECA.  
 DR NON\_TER 1  
 FT NON\_TER 25  
 SQ SEQUENCE 25 AA; 2609 MW; 55BA20167C76F466 CRC64;

Query Match 59.7%; Score 37; DB 5; Length 25;  
 Best Local Similarity 63.6%; Pred. No. 5.7;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KDATFGTHDG 11

Db 10 KDTAAAFATHAG 20

## RESULT 15

Q9YIL9 PRELIMINARY; PRT; 98 AA.  
 AC Q9YIL9;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Hemoglobin IIB (Fragment).  
 GN GB28.  
 OS Chironomus tentans (Midge).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;  
 OC Chironomidae; Chironominae; Chironomus.  
 OX NCBI\_TaxID=7153;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=tentSto;  
 RA Martin J., Guryev V., Bilnov A.;  
 RT "Isolation and Hybridization in the Differentiation of Holarctic  
 RT Camptochironomus (Diptera, Nematocera).";  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF110171; RAD43851.1;  
 DR HSSP; P02229; 1ECA.  
 DR InterPro; IPR002336; Erythruridin.  
 DR InterPro; IPR000971; Globin.  
 DR Pfam; PF00042; globin; 1.  
 DR PRINTS; PR00611; ERYTHCRUORIN.  
 DR PROSITE; PS01033; GLOBIN; 1.  
 KW Heme; Oxygen transport; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 98  
 SQ SEQUENCE 98 AA; 10868 MW; 39126808F80BF20 CRC64;

Query Match 59.7%; Score 37; DB 5; Length 98;  
 Best Local Similarity 63.6%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KDKATFGTHDG 11  
 ||| |||  
 DB 20 KDTAAAFATHAG 30

Search completed: August 28, 2003, 18:37:54  
 Job time : 28.5 secs